AN ONTOLOGICAL REPRESENTATION OF A TAXONOMY FOR CYBERCRIME

Research

Barn, Ravinder, Royal Holloway University of London, UK, r.barn@rhul.ac.uk

Barn, Balbir, Middlesex University, UK, b.barn@mdx.ac.uk

Abstract

The modern phenomenon of cybercrime raises issues and challenges on a scale that has few precedents. A particular central concern is that of establishing clarity about the conceptualization of cybercrime and its growing economic cost to society. A further related concern is focused on developing appropriate legal and policy responses in a context where crime transcends national jurisdictions and physical boundaries. Both are predicated on a better understanding of cybercrime. Efforts at defining and classifying cybercrime by the use of taxonomies to date have largely been descriptive with resulting ambiguities. This paper contributes a semi-formal approach to the development of a taxonomy for cybercrime and offers the conceptual language and accompanying constraints with which to describe cybercrime examples. The approach uses the ontology development platform, Protégé and the Unified Modeling Language (UML) to present an initial taxonomy for cybercrime that goes beyond the descriptive accounts previously offered. The taxonomy is illustrated with examples of cybercrimes both documented in the Protégé toolset and also using UML.

Keywords: Cybercrime, Taxonomy, Ontology, Conceptual Model, Protégé.

1 Introduction

The scale of challenges and issues raised by the phenomenon of cybercrime has few precedents in modern society. A particular central concern is that of establishing clarity about the conceptualization of cybercrime and its growing economic cost to society. A further related concern is focused on developing appropriate legal and policy responses in a context where crime transcends national jurisdictions and physical boundaries. Both are predicated on a better understanding of cybercrime.

In the UK, the estimated cost of £27 billion per year proposed by Detica (2011) has been questioned due to its lack of rigour and transparency despite the recognition that 'modeling cybercrime is a complex and difficult exercise' (Detica 2011:3). More nuanced efforts such as that by Anderson et al (2012) are also perceived to have limitations due to reliability on case studies (McGuire & Dowling, 2013).

One possible factor leading to the difficulties of estimation is the lack of well-formed definitions and classification schemes able to account for the range of cybercrimes. Consequently, this area continues to be a recipient of significant research effort that aims to resolve the ambiguity and degree of contextual mutability around notions of cybercrime (Fafinski et al, 2010, Donalds & Osei-Bryson, 2014).

This paper aims to address the lack of well-formed definitions by contributing a semi-formal model that is able to classify cybercrimes within a taxonomy that has an ontological foundation. Given the dynamic nature of technology, the proposed taxonomy is developed both as a conceptual model and as an ontology that can be extended both in terms of new concepts and new types of cybercrime. The on-

tology benefits from the use of standard tools and methods to develop ontologies and will become publicly available. This then, is an initial step towards a reference model for cybercrime. The research presented builds on a preliminary model produced as part of the CYBEROAD project for the EU Seventh Framework Programme (CyberRoad, 2015).

The remainder of the paper is structured as follows: Section 2 introduces the background to cybercrime. Component elements of cybercrime are discussed, together with challenges associated with defining cybercrime. Existing efforts at classifying cybercrimes are reviewed. Section 3 provides introductory material on the underlying technological foundations we use for developing our proposed taxonomy. Core technologies such as the Unified Modelling Language (UML), Ontology Web Language (OWL) and the widely used ontology development platform, Protégé from Stanford University (Protégé, 2015) are described. Section 4, outlines the key stages of our approach to developing the taxonomy. In particular the relationship between conceptual models produced in UML and Ontologies is delineated. The method is differentiated from other recent efforts to define approaches to developing taxonomies for cybercrime. In particular, we demonstrate difference to efforts by Nickerson et al. (2013) and Land et al. (2015). Section 5 presents the main contribution of our work in the form of a taxonomy for cybercrime that has capability for both extension with new concepts and classifying new cybercrimes. The taxonomy is presented as a UML conceptual model and as a Protégé ontology. We recognize that our taxonomy is a case of emergent theory building in the sense of Doty & Glick (1994) and so we present an initial evaluation using two case study scenarios in Section 6. Finally, in section 7, we present concluding remarks and further research plans for extending the ontology as a reference model for cybercrime.

2 Background to Cybercrime

Cybercrime, like traditional crime, can be conceptualised as historically and culturally situated, and encapsulated within a social and political ideological framework (Christie, 2004). This creates a challenge for an efficient definition of cybercrime and also for classification or categorisation of cybercrimes. The latter is particularly important as it enables key use cases such as better costing of cybercrime to be developed. Syntheses of categorisations proposed by some scholars suggest cybercrimes can be understood in three principal ways (Oates, 2001; Wall, 2005; European Union, 2007; Anderson et al, 2013). These are (1) traditional crimes that are contingent on the use of technology, (2) publication of illegal content on the internet, and (3) crimes that occur within technological forums. The UK Home Office, however, has expressed a preference for opting for a two-fold categorisation: 'Cyber-dependent' and Cyber-enabled' crimes (McGuire & Dowling, 2013). Here 'cyber-dependent' encompasses 'new' crimes made possible by technology such as malware, hacking, and viruses, while 'cyber-enabled' refers to 'old' crimes such as theft, fraud, and harassment which are committed using computers. Gordon and Ford (2006) also propose a two-fold dichotomy, but they suggest that crimes are 'techno-centric' (Type 1) or 'people-centric' (Type-2). The first category is aligned with cyber-dependent crimes, the latter includes a strong social engineering context. Again there is an element of some role being played by technology. Ngafeeson (2010) whilst, acknowledging the role technology as an enabler or a dependent variable, proposes that cybercrime also needs to be understood from a sociological perspective and so draws upon crime theory to suggest a classification scheme based on understanding the motivations of perpetrators. Kshetri (2006) discusses motivation as intrinsic or extrinsic, where the former can have a superior impact to the latter. Others, such as Yar (2006), have subdivided cybercrime into areas of harmful activity that illustrate a range of activities and behaviours rather than focussing on specific offences. Thus this is an attempt to focus on the impact of a cybercrime.

A common aspect of all these approaches, is that that the classifications present a relatively simple form of theory building for explanation – the use of typologies (Doty & Glick, 1994). One effort at formalising cybercrime more precisely is that by Stabek et al. (2010) who attempt to define a set-theoretic view of cyberscams.

A further issue is that of the legal viewpoint. As Fafinski et al. (2010) note: "Cybercrime is not a legal term of art". An implication of this is that prefix of cyber raises legal questions and may even fall outside of criminal justice processes such as the case of cyber-rape in Second Life (http://secondlife.com). Similarly, the destruction of underwater data cable may or may not be classified as cybercrime. Within their discussion, they propose a variant on a three fold typology, based on that by Wall (2005), that categorises cybercrime along three dimensions: crimes that are against technology (such as unauthorised access), crimes that use technology (such as fraud), or crimes in the machine that are content related (such as illegal trading of sexual materials). The latter can be debated at much length, for example, content stored on a computer is surely *using* technology?

What is clear is that the various classifications focus on different facets of cybercrime and a significant shortcoming is that there is no existing classification that presents an integrated view of cybercrime that adequately conceptualises cybercrime as involving: "a number of key elements and questions that include *where* the criminal acts exist in the real and digital worlds (and *what* technologies are involved in carrying out the crimes), *why* are malicious activities initiated and *who* is involved in carrying out the malicious acts?" (Finklea & Theohary, 2012:2).

A key question is thus how to present a synthesis of these different classifications in a form that can accommodate all these classifications and potentially allow new inferred classifications to emerge as the new taxonomy is utilized. Given the dynamic nature of technology and the emergence of new forms of cybercrimes, this is crucial. We argue that the global, informational, and distributed nature of the internet adds to the complexity of the phenomenon (Wall 2015). The compression of time, space, and distance allows not only for easy access to potential targets of cybercrime, but also possibilities of large-scale cybercriminal activity. By proposing a taxonomy of cybercrime, this paper contributes to the literature to help explore the academic value of this elusive and contested concept. In the remainder of the paper, we lay out the foundational technology for a proposal for an integrated taxonomy for classifying cybercrimes.

3 Ontologies and Unified Modeling Language (UML)

Earlier, we noted that efforts at classifying cybercrimes have largely offered descriptive accounts of how a cybercrime might be classified. Little attempt has been made to produce more formal classifications. In this section, we provide a short description of how core technologies of ontology development and UML can be used as foundational form for classifying cybercrime.

Ontologies are used to capture knowledge about some domain of interest. Ontology describes the concepts in the domain and also the relationships that hold between those concepts. A widely cited definition of ontology is that offered by Gruber (1993:199): "An Ontology is an explicit specification of a conceptualization." This has been variously refined to include additional characteristics such as 'formal' and 'shared', where the former is implying machine readability and the latter indicates that the conceptualization is an acceptance by a community. Ontologies are content theories about the sorts of objects, properties of objects and relations between objects possible in a specified knowledge domain (Chandrasekeran et al., 1999) and as such enables the study of "…the most pervasive features of reality" (Fettke & Loos, 2003:2945).

Developing an ontology for a given domain provides several advantages arising from the formal underpinnings, that is, first order logics. There is provision for a well-defined glossary for a domain; real-world semantics; the identification of inappropriate constructs revealing problems in the definition, interpretation and/or usage of concepts (through built-in reasoners) and support for evolutionary development (and inference) of new knowledge relating to a domain.

There are several ontology representation languages available. This paper uses the W3C recommendation ontology language OWL (Web Ontology language) in its version OWL2 to represent the ontology (http://www.w3.org/2001/sw/wiki/OWL). An OWL ontology consists of Individuals, Properties and Classes. Individuals, represent objects in a given domain. Properties are binary relations on individuals

and are used to link two individuals together. Properties can be inverse to allow a link between two individuals to be traversed in both directions. OWL classes are interpreted as sets that contain individuals. Classes can be organized into taxonomies of super-class/sub-class hierarchies and are described formally using logics.

The Unified Modelling Language (UML) is an OMG standard used for describing software intensive systems using diagrammatic representations (OMG, 2015). This paper uses a subset of UML (the class diagram and object diagrams) as a conceptual modeling technology for representing our proposed cybercrime taxonomy. The relationship between UML conceptual models and ontologies has been the subject of much research and there several examples that show how the two approaches can be integrated (Wang & Chan, 2001; Paulheim et al., 2011). Based on this, the following mapping between the two technologies is shown in Table 1 below.

UML Concept	OWL2 DL / Protégé Concepts
UML Class	Class
UML Attribute	ObjectProperty
UML Association End	ObjectProperty
UML Association	ObjectProperty

Table 1 Mapping UML model elements to OWL2 DL

4 Methodology

Various methods exist for developing taxonomies and as Nickerson et al. (2013) point out, they also tend to be ad hoc. Additionally, most approaches also vary in terms of formality, rigour and evaluation (Land et al. 2015). Recent efforts by Nickerson et al. and more latterly, Land et al. have both attempted to lay out prescribed steps for building a taxonomy that begin from either an inductive approach (empirical to conceptual) or a deductive approach (conceptual to empirical). Partly, this division appears to be explained by discipline dependent viewpoints. Thus they note that: a "typology" (common in the social sciences) is drawn from conceptual classification, while a "taxonomy" (widely used in the biological sciences) is drawn from empirical classification (Bailey, 1994). Integration between these two viewpoints occurs in the latter stage when there is a "working taxonomy" (Figure 2 in Land et al. 2015). We also note that both approaches do not specifically comment on the working language or format of the taxonomy. This latter point is important if we want to utilize taxonomies through systems and therefore require that there is a minimum machine manipulatability aspect to the approach.

Given these limitations of delayed integration of conceptual with empirical evidence and the lack of a working language for a taxonomy, we adopt a methodology that incorporates the following features:

- Closer and earlier integration between deductive and inductive approaches to developing the meta features of a taxonomy.
- Utilisation of established language constructs for specifying the conceptual features of the taxonomy (based on UML 2.0 standard).
- A design science approach that is in the spirit of "how to do" proposed by March & Smith (1995) that is demonstrated by the use of an ontology published using the XML/OWL language (http://www.w3.org/2001/sw/wiki/OWL).
- Support for open world analysis that allows the taxonomy to act as a reference model that can easily be extended both in terms of meta concepts and taxonomic instances.
- Support for machine readability and hence interaction with other systems.

The Cybercrime Taxonomy proposed in this paper in section 5 uses the methodological framework shown in Figure 1. There are three stages in this process. We describe each stage below.

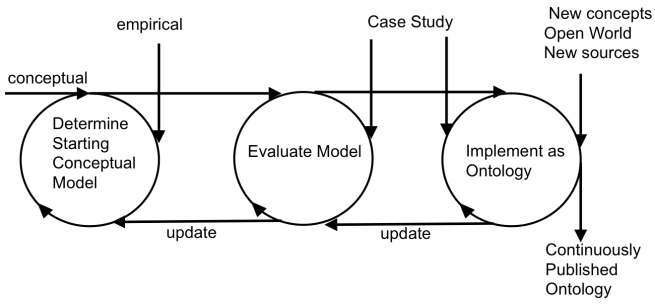


Figure 1 Taxonomy development method

Determine Starting Conceptual Model

As in the methodology originally proposed by Nickerson et al. (2013) and subsequently extended by Land et al. (2015), the first step is to identify the meta characteristics required for the objective of the taxonomy. These characteristics are derived from both academic literature including victimization surveys, and empirical case studies documented in the public domain such as cybercrimes reported in newspapers (CyberRoad 2015). In contrast to prior approaches, we document the meta characteristics using a conceptual modeling approach with the language tooling offered by the UML 2.0 standard (OMG, 2015). A conceptual model is the first artifact of that activity. We suggest that this adds significant rigour to other approaches where no such formal language is used.

Evaluate Model

This stage is closely linked with the previous one and its main purpose is to refine the initial conceptual model with further case study examples which are used to evaluate the model in terms of the dimensions proposed by Moody (2005) such as completeness and consistency. Evaluation is enacted through the development of object models. The benefit of using object models is the possibility of ensuring a well-defined semantics. Here, we limit the semantics of this model to be a collection of object models that are instances of the semantic model. The semantic model comprises objects and slots that contain values. Additionally, there are well-formed rules that determine how an instance model is deemed to be correct with respect to the conceptual model. As new case studies end up being represented by object models, new properties, types and relationships are added to the starting conceptual model indicated by the *update* arrow.

Implement as Ontology

The updated conceptual model represents a closed world. For a cybercrime taxonomy to deliver the benefits outlined in the introduction section, it needs to be constantly updated and evolved to meet the needs of multiple stakeholders. The border-free aspect of cybercrime places demands on the need for efforts at standardisation and so a further requirement is a taxonomy that can also act as a reference model for the domain of cybercrime. For both these requirements, domain ontologies can be used to represent a standardised model of the cybercrime domain (van Heijst et al, 1997). Consequently, this step translates the conceptual model developed from the previous stages to an ontology-based representation using Protégé. Established mechanisms for translating UML concepts to OWL/Protégé concepts are used (Table 1). The resulting ontology is described in the next section along with an evaluation of the ontology using a number of perceived cyber-crimes.

5 A Taxonomy for Cybercrime

In this section we present our proposal using an ontological basis for describing taxonomy for cybercrime by noting two observations from recent research. Firstly, Wand & Weber (2004) note that theories of ontology sometimes function like taxonomies. Secondly, as noted by Nickerson (2013), a taxonomy may be a step toward a future ontology. We will argue that this perceived inter-dependency is advantageous in discussions of classifications of cybercrimes as it allows for evolution and debate in classifications.

Following the cyclical nature of the method outlined in the earlier section, we show how the conceptual model (using UML class diagrams) and its translation into OWL2 DL using the open source ontology tool Protégé is closely integrated.

5.1 Cybercrime Conceptual Model

Using the existing literature, we propose a conceptual model for cybercrime. The model is shown in its entirety in Figure 2. We refer to the model in the following discussion.

As noted earlier, the model concepts were derived from existing efforts to classify cybercrimes. At the most abstract level, we are able to state that: An *Agent* is motivated by either an *Intrinsic* desire or an *Extrinsic* need to commit an *Action*. *Actions* are perceived as *Crimes* depending upon a receiver *Agent's Viewpoint*. If an *Action* is a *Cybercrime* (and so subsuming the concept of *Crime*), then the *Cybercrime* must be mediated through a *TechnologyRole*. That is, some form of technology must be involved to enact a Cybercrime. An *Action* must have a *Target* and there must be some type of *Impact* endured by the Target. An *Impact* is the effect of a crime on a Target and can be *Economic*, *Psychological* or *Geo-political*. A Crime is an *Action* that is a *TraditionalCrime* or a *CyberCrime*.

In discussions of traditional crime, criminological explanations are nuanced and range from conceptualizing crime as a 'rational' act, to others that locate criminal activity within the paradigms of state power and control, social construction, and a product of the constraints of a social and economic environment (Stucky & Krohn, 2009). The cybercrime literature is largely under-theorized at present but there have been attempts to apply some existing traditional crime theories to crimes committed in cyberspace (Yar, 2005; Li, Zhang & Sarathy, 2010). Such theories discuss the conditions such as 'opportunity', 'availability of potential target/technology', 'absence of legal guardians', and 'potential gain' as a motivator to cybercrime. Notably, motivation as the driving force behind cybercrime has become the key focus. Some of the earliest efforts to explain motivations for cybercrime originated in the work of Furnell (2001) and naturally builds on criminological literature. Hence, the Motivation model element exists to account for both Furnell's explanations of motivations and those proposed by Ngafeeson (2010). Cybercrimes including defrauding, theft, piracy, hacking, and viruses may be conceptualized within the motivation paradigm. Such motivation maybe intrinsic or extrinsic within a moral, ideological, psychological and financial framework. Elements of power and control can be seen as inherent within such motivations. Additionally, the *Agents*, the *Targets*, and the victims of cybercrime may be *Individuals*, Organisations or States who operate within such motivational arenas. *Targets* can also be *TargetTechnology* to account for *Cybercrimes* committed against the computer (Wall, 2005).

The taxonomy allows for crimes to be described as a crime (or cybercrime) or just a normative action through the use of a *Viewpoint* construct. *Viewpoints* are possessed by *Observers* of *Actions*. Many of the existing efforts to either construct taxonomies for cybercrime or to classify cybercrimes refer to the centrality of the role played by technology. A frequently cited effort at classification is that by Gordon & Ford (2006). Others include classifications by Wall (2005), Jahankhani et al. (2014) and Yar (2006). The latter introduced state actors as both initiators and targets of cybercrime. As mentioned above, much of the role of technology in cybercrime has now been synthesized and is perhaps most succinctly captured by a UK Home Office report that described two broad categories of cybercrime: cyber-dependent crimes and cyber-enabled crimes (McGuire & Dowling, 2013). In our conceptual model, The *TechnologyRole* concept has two (disjoint) subtypes: *Enabled* and *Contingent* that describes this

classification. One feature of cybercrime is the frequency of contact and the use of non-physical locations (i.e. cyberspace) where that contact occurs (Gordon & Ford, 2006). This concept and its relationship constraint is presented in the conceptual model.

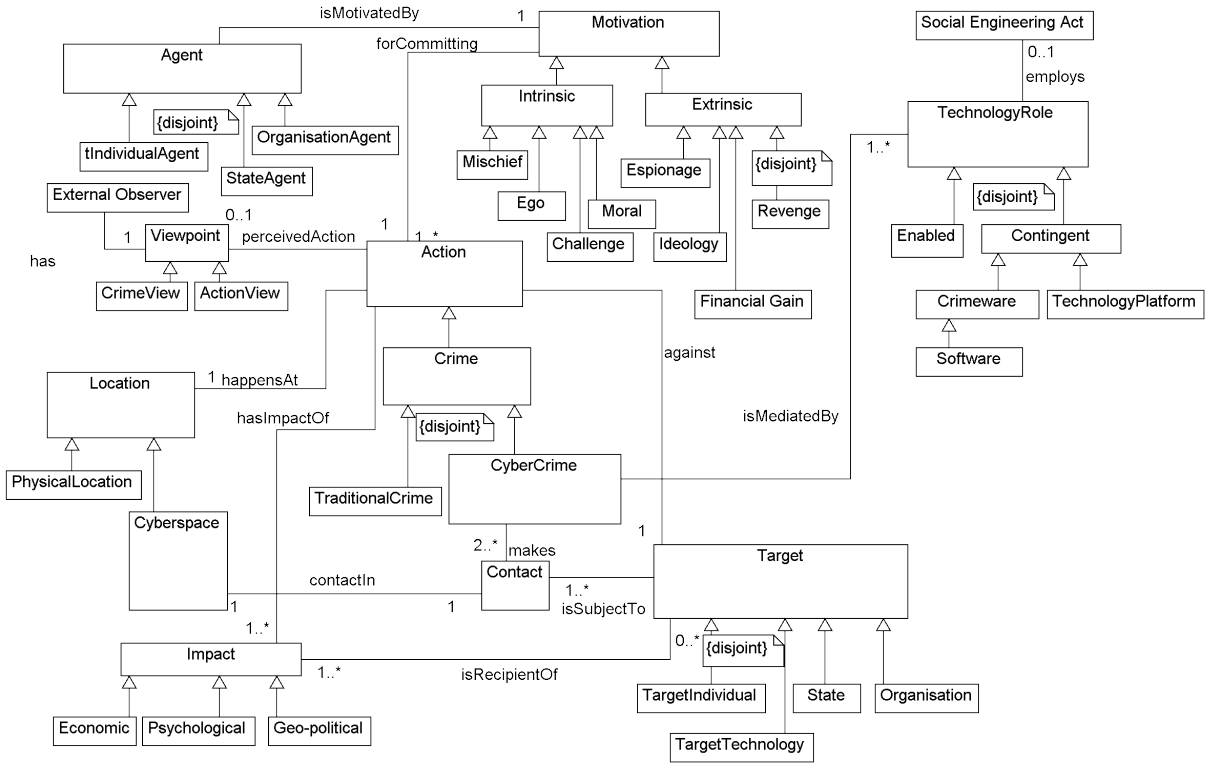


Figure 2 Cybercrime conceptual model

Within the model there are several class hierarchies that are of interest. An Agent can be disjoint with *IndividualAgent*, *StateAgent* or *OrganisationAgent* representing different levels of Agent engagement committing some *Action*. A *Target* has a similar subtyping hierarchy: *TargetIndividual*, *State*, *Organisation* and *TargetTechnology* (to account for infrastructure).

While the Mcguire & Dowling (2013) classification offers a straight-forward bifurcation of cybercrime, there are further subtleties to categorise. Hence we introduce a further subtyping (based on Gordon & Ford) of *Contingent* to include the notion of *Crimeware* (software used in the commission of a criminal act whose use is not involuntary) as opposed to *TechnologyPlatform* that is merely a platform whose contingency enables a crime. For example, a security flaw in a Web Browser is not crimeware, but the flaw makes possible for a cybercrime to be committed. Similarly, TechnologyPlatform also allows for the use of technology to be container for content (akin, to Wall's "Crime in the Machine"). A further consideration is the role of social engineering acts to manipulate targets to enable a Cybercrime. We account for this by the *employs* association relationship between *TechnologyRole* and *SocialEngineeringAct*.

5.2 Constraints

A conceptual model that offers only a visual representation of cybercrime is not sufficient, as there a number of constraints that are not immediately clear to the interested reader. We offer some sample

constraints using the OMG standard Object Constraint Language (Warmer & Kleppe, 2003) to illustrate this need and hence the benefits of moving to an ontological foundations for a taxonomy.

5.2.1 Abstract Classes

We propose that some classes are abstract. That is, there are no instances of that class. Such classes include: `TechnologyRole, Motivation, Location, Target, Impact`, and `Agent`. We illustrate this constraint for TechnologyRole and Agent as:

`TechnologyRole`

```
TechnologyRole.allInstances->Select(oclType=TechnologyRole)isEmpty
/* The set of all instances of TechnologyRole is empty */
```

`Agent`

```
Agent.allInstances->Select(oclType=Agent)isEmpty
/* The set of all instances of Agent is empty */
```

5.2.2 Loop Invariants

In many conceptual models, there is a requirement to ensure that navigation of associations results in returning to the instance of a class from which the navigation was initiated. In the model on Figure 2, there is one such invariant.

We want to assert that the target of a crime results in an impact to the target. That is, when a crime is committed against a target, there is an impact of the crime on that target. We express the constraint as follows:

`Target`

```
self.action.impact.target=self
```

5.2.3 Perceptions of Crime

An external observer may have a perception that an action is a crime or may simple be an action in the most general sense. We account for this by adding constraints on the association relationship between Viewpoint subclass hierarchy and Action.

`CrimeView`

```
self.action->forAll(oclIsKindOf(Crime))
```

`Action View`

```
self.action->forAll(oclIsKindOf(Action))
```

5.2.4 Cybercrime

OCL constraints allows us to be more precise about what is a cybercrime. We assert that a cybercrime is mediated by a TechnologyRole and targets either Individuals, States, Organisations or TargetTechnology and all the Contact arising from the crime is made in Cyberspace. This is expressed as a constraint in the following manner:

`Cybercrime`

```
self.technologyrole->notEmpty ∧ self.contact->forAll(c:Contact |
(c.cyberspace->notEmpty))
```

5.3 Ontological Representation of Cybercrime

There are several advantages in using ontologies to represent theories based on taxonomic structures. Ontologies offer parsimony of concepts, are easily extendable, can be explanatory and more importantly can present hidden inferences because of the formal underpinnings. In this section we present two aspects of the ontology, the taxonomic (hierarchical class structures) and the relationships between classes. through an implementation of bi-directional OWL object properties. The taxonomic structure is a straightforward translation into OWL classes and is shown in Figure 3.

Figure 3 Taxonomy class structures in OWL/Protégé

Experience from ontology engineering shows that inverse relationships should be included in an ontology in order to fully specify concept relationships (Lantow & Sandkuhl, 2015). The association relationships are implemented as bi-directional OWL properties and the details are shown in Table 2 and Figure 4 below.

Domain	Object Property name / Inverse	Restriction	Range
Action	against / subjectedAction	Some / Min 1	Target
Impact	causedBy / hasImpactOf	Some / Min 1	Action
Social Engineering Act	employedBy / employsSocialEngineering	Some / Min 1	TechnologyRole
Location	For / happensAt	Some / Min 1	Action
Motivation	forCommitting / motivatedBy	Min 1 / Some	Action
Motivation	motivationFor /	Some /	Agent

	isMotivatedBy	Min 1	
Impact	On / recipientOf	Min 1 / Min 1	Target
Viewpoint	perceivedBy / perceives	Some / Exactly 1	External Observer
Viewpoint	perceptionOf / perceivedFrom	Exactly 1 / Some	Action
Cybercrime	isMediatedBy / mediates	Min 1 / Some	TechnologyRole
Contact	usedBy / makes	Some / Min 2	Cybercrime
Contact	contactIn / locationFor	Min 1 / Min 1	Cyberspace

Table 2 Conceptual relationships documented in OWL

To illustrate how this works, the following diagram shows a model snippet of an associations between classes and the equivalent highlighted OWL object properties.

Figure 4 Model snippet of association relationships and their implementation in OWL/Protégé

6 Evaluation

We outline two initial evaluations of our model for cybercrime. First, we use the formal ontology documented using OWL/Protégé to describe two examples of well known cybercrimes: the Nigerian 419 scam used as an exemplar by Stabek et al. (2009) and the more recent CryptoLocker malware example. Second, we use the UML conceptual model to explore the Snowden revelations of 2013 (Landau, 2013) to illustrate social constructions of crime dependent upon an external observer's viewpoint.

6.1 Nigerian 419 Scam and CryptoLocker

The Nigerian scam is described by Land et al. (2015:6) as: "an unsolicited email (or fax or any other methods of delivery via the internet) detailing an unfortunate story of the sender (i.e. scammer) and (s)he has a fortune, but (s)he needs the victim to supply an overseas bank account to transfer the money or request a small amount of money for a short term. In return for the victim's assistance, the scammer promises to share some of the fortune with the victim, say 10% for the total amount of wealth".

The Cryptolocker malware is software that encrypts various files on the user's computer and demands the owner to pay the malware authors in order to decrypt the files (Mustaca, 2014).

For both these well-known examples, we use the taxonomy structure in the ontology model in Protégé and the supporting conceptual model together to classify and store classifications of cybercrimes. To do so, the ontology designer can ask the following series of questions and create instances (OWL Individuals) of the appropriate class. Table 3 illustrates the questions and instances created for both examples of cybercrime.

Method Question	**Nigerian 419**	**CryptoLocker**
What is the type of Agent?	Individual/Organisation	Individual/Organisation
What is the Motivation for committing the action	Financial Gain	Financial Gain
What is the perceived nature of the action? Is it a crime?	Crime	Crime
Is it mediated by Technology?	Yes	Yes
What role is technology playing?	Technology Platform	Software
Is social engineering employed	Social Engineering Act	Social Engineering Act
How is contact being made via Cyberspace	Email	Email
What is the type of Target	Individual	Individual
What is the Impact	Economic / Psychological	Economic / Psychological

3 Methods for differentiating between cybercrimes

Table 3 provides a summary of the key concepts. However, the documentation of the cybercrime in the Protégé toolset can show much more detail, including, critically, the relationships between object instances. Figure 5 shows a cybercrime instance – Nigerian-419-scam and its relationships to other objects.

More examples of cybercrimes classified by a very prelimary version of this taxonomy is published in a technical report available from the EU Seventh Framework Programme CyberRoad project (Cyberroad, 2015).

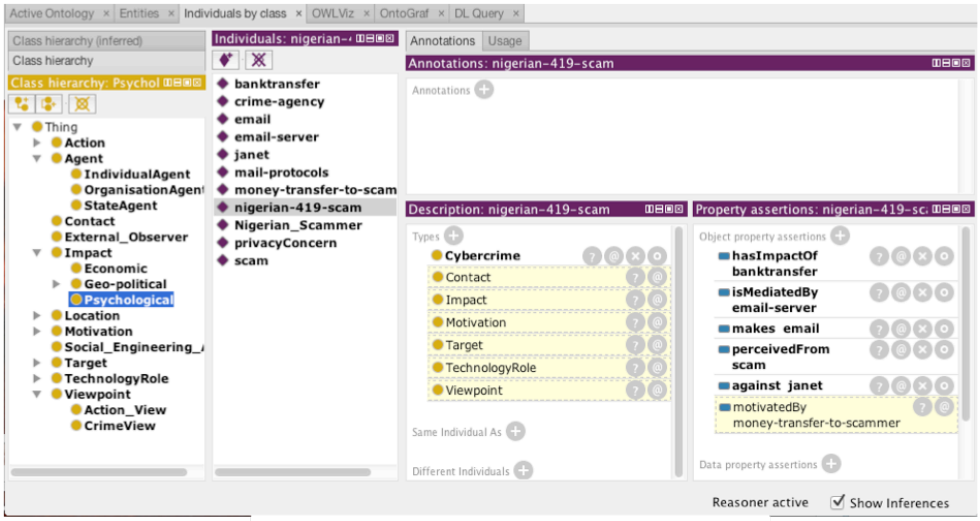
6.2 Snowden Revelations 2013

In June 2013, the Guardian Newspaper in the UK began publishing a series of exposés on bulk data collection conducted by the US National Security Agency. This information was obtained from a cache of classified data obtained illegally by Edward Snowden who was working as a contractor for Booz Allen Hamilton at the NSA headquarters. Information revealed included several programmes of work that targeted Internet Communications and stored data of "non-US persons". Other leaks included details of the US government spying on Chinese computers and reports that Britain was also conducting massive intercepts of domestic communications (Landau, 2013).

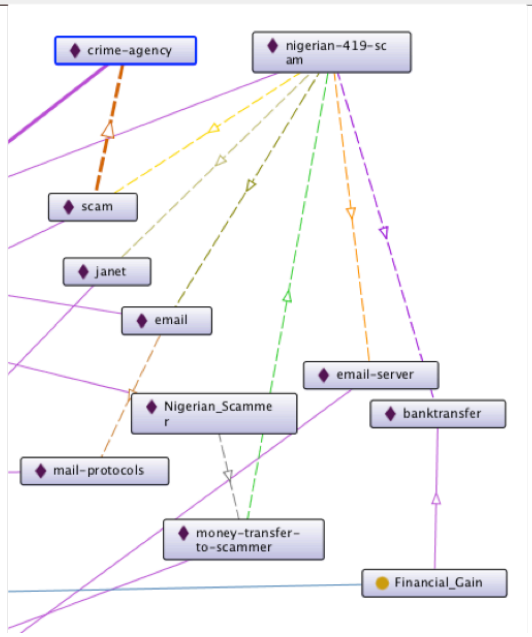
Here, we are not debating the ethical concerns of the actions of Snowden, instead, we wish to show two hypothetical models of interpretation of Snowden's actions using our taxonomic model. The diagram below illustrates the two perspectives. In the upper part of Figure 6, Snowden is an Individual Agent who is motivated by a Moral concern that presents a reason for his Action. A 'privacy organisation' (External Observer) views Snowden's Action as not a crime hence there are no further instances of the model.

The lower diagram in Figure 6 shows a more complete object model, where the external observer such as a security organisation perceives the Action as a crime and as the crime is contingent on technology, has impacts on the target (state) that are geo-political and (some) contact was made in cyberspace,

the crime is an example of a cybercime. We can use the invariant presented in 5.14 to check the valid semantics of both these two object model snippets. The invariant fails for the moral crusader model but holds for the cyber-criminal model.



Exploring in Protégé



Visalisation

Figure 5 Nigerian 419 scam in OWL/Protégé

How does this taxonomy compare with other established models of cybercrime? The widely cited typology for cybercrime by Gordon & Ford (2006) would classify the Nigerian 419 Scam as a Type II cybercrime in that it is facilitated by programmes that are not seen as crimeware and there are repeated contacts with the target user (Gordon & Ford, 2006:16). CryptoLocker on the other would be seen as Type I cybercrime as it has been facilitated by the introduction of crimeware programmes. The Snowden example however, would not be classified as a cybercrime. The Gordon & Ford classification does not allow us to see the wider context of the motivation of the perpetrator, nor the nature of the impact on the target. It also does not allow us to consider sociological aspects of the crime with respect to the perspective from which the cybercrime is being classified. These additional details, we contend, are beneficial if we are to develop more effective models for costing cybercrime or for developing appropriate policies and education around cybercrime.

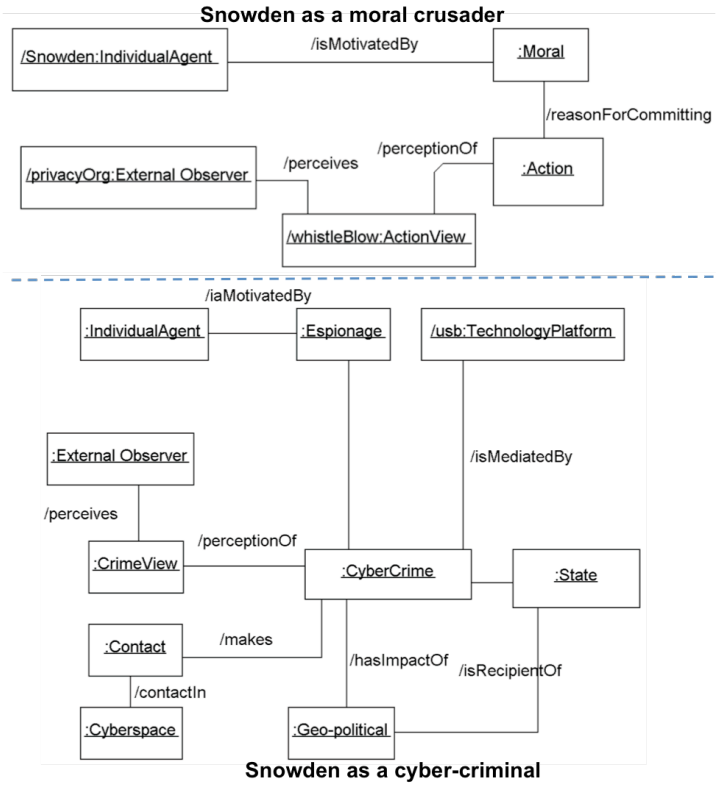


Figure 6 Object model snippets for Snowden example

6.3 Validity Concerns

The production of IT artifacts in design science raise concerns of instantiation validity that are analagous to the concepts of construct validity in survey research (Straub et al., 2004, Lukyanenko, 2014). However, there are further challenges. Firstly, realistic IT artifacts are more complex and expensive to build, secondly the artifacts occupy a large design/instantiation space. For example, there may be multiple ways to code a particular algorithm. Further, emergent properties may arise that are difficult to predict. Both the UML model and its ontology representation are examples of a mid-range theory, that is, they are moderately abstract (i.e., they do not purport to explain everything) but are "close enough to observed data to be incorporated in propositions that permit empirical testing" (Merton, 1949). We have illustrated the validity of these models but full exploration of the design space requires further empirical evidence by independent scholars. To this end, we have published the ontology for further scrutiny at http://goo.gl/uKmYL1.

7 Concluding Remarks

In this paper, we have examined existing definitions of cybercrime and like Gordon & Ford (2006), we conclude that there is a lack of clarity of common usage and classification of cybercrime. This paper addresses this concern by advancing a new taxonomy of cybercrime that has semi-formal rigour in that the taxonomy is based on a set of concepts described through the use of a UML model and accompanying wellformedness rules. To account for the open world nature of cybercrime, that is, we cannot foresee future types of cybercrimes, we have proposed an ontology for cybercrime that has been implemented in OWL/Protégé and has been made publicly available through the Web Protégé environ-

ment. This ontology is open to extension, both in terms of new concepts and new examples of cybercrime.

Our taxonomy is a first step in developing both a well-formed taxonomy and the surrounding context of a cybercrime classification that will result in better understanding to enable the development of more effective models for costing cybercrime. Our future work will include developing new axioms to support reasoning about cybercrime in order to ask specific questions such as cost, and efficacy of training to address specific cybercrime concerns.

References

Almeida, J. P., Cardoso, E. C. S., & Guizzardi, G. (2010, October). On the goal domain in the RM-ODP Enterprise Language: An initial appraisal based on a foundational ontology. In *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2010 14th IEEE International* (pp. 382-390). IEEE.

Lantow, B., & Sandkuhl, K. (2015). From Visual Language to Ontology Representation: Using OWL for Transitivity Analysis in 4EM. In Practice of Enterprise Modelling, 2015, http://ceur-ws.org/Vol-1497/PoEM2015.

CyberRoad. (2015). *Project*. Retrieved 19 November 2015, from http://www.cyberroad-project.eu/en/project/.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them?. *IEEE Intelligent systems*, (1), 20-26.

Christie, N. (2004). *A suitable amount of crime*. Psychology Press.

Donalds, C., & Osei-Bryson, K. M. (2014). A Cybercrime Taxonomy: Case of the Jamaican Jurisdiction.

Doty, D. H., & Glick, W. H. (1994). Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of management review*, *19*(2), 230-251.

Gordon, S. & Ford, R. (2006). On the Definition and classification of Cybercrime. *Journal of Computer Virology*. 2:13-20.

Fettke, P., & Loos, P. (2003). Ontological evaluation of reference models using the Bunge-Wand-Weber model. *AMCIS 2003 Proceedings*, 384.

Finklea, K. M., & Theohary, C. A. (2012, May). Cybercrime: conceptual issues for congress and US law enforcement. Congressional Research Service, Library of Congress.

Furnell, S. M. (2001). The problem of categorising cybercrime and cybercriminals. In *2nd Australian Information Warfare and Security Conference*.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, *5*(2), 199-220.

Jahankhani, H., Al Nemrat, A., & Hosseinian-Far, (2014). Cybercrime classification and characteristics. In: *Cybercrime and Cyber Terrorism Investigators' Handbook*. Syngress.

Kshetri, N. (2006) The simple economics of cybercrimes. Security & Privacy, IEEE, 4, 1, p. 33-39.

Land, L., Smith, S., & Pang, V. (2013, June). Building a Taxonomy for Cybercrimes. In *PACIS* (p. 109).

Landau, S. (2013). Making sense from Snowden: What's significant in the NSA surveillance revelations. *IEEE Security & Privacy*, (4), 54-63.

Li, H., Zhang, J., & Sarathy, R. (2010). Understanding compliance with internet use policy from the perspective of rational choice theory. *Decision Support Systems*, *48*(4), 635-645.

Lukyanenko, R., Evermann, J., & Parsons, J. (2014). Instantiation validity in IS design research. In *Advancing the Impact of Design Science: Moving from Theory to Practice* (pp. 321-328). Springer International Publishing.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision support systems*, *15*(4), 251-266.

Merton, R. (1949). On Sociological Theories of the Middle Range. In: Merton, R. (ed.) Social Theory and Social Structure, pp. 39–53. The Free Press, New York (1949)
Mustaca, S. (2014). Are your IT professionals prepared for the challenges to come?. *Computer Fraud & Security*, *2014*(3), 18-20.
McGuire, M., & Dowling, S. (2013). Cyber crime: A review of the evidence. *Summary of key findings and implications. Home Office Research report*, *75*.
Moody, D. L. (2005). Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. *Data & Knowledge Engineering*, *55*(3), 243-276.
Ngafeeson, M. (2010). Cybercrime Classification: A Motivational Model. *2010 Southwest Decision Sciences Institute Conference.*). Retrieved 13 November 2015, from http://goo.gl/VzY9MN.
Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336-359.
OMG. (2015). *UML 2.5*. Retrieved 19 November 2015, from http://www.omg.org/spec/UML/2.5/PDF/
Paulheim, H., Plendl, R., Probst, F., & Oberle, D. (2011, April). Mapping pragmatic class models to reference ontologies. In *Data Engineering Workshops (ICDEW), 2011 IEEE 27th International Conference on* (pp. 200-205). IEEE.
Protégé. (2015). Stanford Center for Biomedical Informatics Research, The Protégé ontology editor and knowledge acquisition system, http://protege.stanford.edu/.
van Heijst, G., Schreiber, A., Wielinga, BJ. (1997) Using explicit ontologies in KBS development. International Journal of Human-Computer Studies 45:183-292.
Stabek, A., Brown, S., & Watters, P. (2009, July). The Case for a Consistent Cyberscam Classification Framework (CCCF). In *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC'09. Symposia and Workshops on* (pp. 525-530). IEEE.
Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *The Communications of the Association for Information Systems*, *13*(1), 63.
Stucky, T. D., & Krohn, M. D. (2009). *Researching theories of crime and deviance*. Oxford University Press, USA.
Wall, D.S. (2005). The internet as a conduit for criminal activity. In: Pattavina, A. (Ed.), Information Technology and the Criminal Justice System. Sage Publications, USA, ISBN 0-7619-3019-1.
Wall, D. S. (2015). Dis-Organised Crime: Towards a Distributed Model of the Organization of Cybercrime. *The European Review of Organised Crime*, *2*(2).
Wang, X., & Chan, C. W. (2001). Ontology modeling using UML. In *OOIS 2001* (pp. 59-68). Springer London.
Warmer, J. & Kleppe, A. (2003). The Object Constraint Language, Second Edition. Getting your Model ready for MDA, Addison Wesley.
Yar, M. (2005). The novelty of 'cybercrime'an assessment in light of routine activity theory. *European Journal of Criminology*, *2*(4), 407-427.
Yar, M. (2006). Cybercrime and Society. Sage Publication Ltd, London.